Deutsches Institut für
Wirtschaftsforschung

DIW BERLIN

Discussion Papers

962

Maria Nieswand • Astrid Cullmann •
Anne Neumann

Overcoming Data Limitations in Nonparametric Benchmarking: Applying PCA-DEA to Natural Gas Transmission

Berlin, December 2009

IMPRESSUM

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin
Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

Available for free downloading from the DIW Berlin website.

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN.
Papers can be downloaded free of charge from the following websites:

http://www.diw.de/english/products/publications/discussion_papers/27539.html
http://ideas.repec.org/s/diw/diwwpp.html
http://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalbrowse&journal_id=1079991

Overcoming Data Limitations in Nonparametric Benchmarking: Applying PCA-DEA to Natural Gas Transmission

Maria Nieswand* **Astrid Cullmann** **Anne Neumann**
DIW Berlin, Mohrenstrasse 58, D-10117 Berlin, Germany

December 2009

Abstract

This paper provides an empirical demonstration for a practical approach of efficiency evaluation against the background of limited data availability in some regulated industries. Here, traditional DEA may result in a lack of discriminatory power when high numbers of variables but only limited observations are available. We apply PCA-DEA for radial efficiency measurement to US natural gas transmission companies in 2007. This allows us to reduce dimensions of the optimization problem while maintaining most of the variation in the original data. Our results suggest that the PCA-DEA methodology reduces the probability of over-estimation of the individual firm-specific performance. It also allows for a large number of original variables without substantially reducing the discriminatory power of the model.

JEL-Codes: C14, L51, L95
Keywords: Efficiency analysis, PCA-DEA, network regulation, natural gas transmission

*Correspondence: email: mnieswand@diw.de, phone: +49 30 89789224, fax: +49 30 89789104.
This paper is a result of the research program on 'Efficiency Analysis', run jointly by DIW Berlin and the Chair of Energy Economics and Public Sector Management at TU Dresden. Earlier versions have been presented at the XI EWEPA in Pisa, the 8th INFRADAY in Berlin, and the Workshop on Exploring Research Frontiers in Contemporary Statistics and Econometrics in Louvain-la-Neuve. We thank Nicole Adler, Victor Podinovski, David Saal, Christian von Hirschhausen and the workshop participants for valuable input and fruitful discussions. The usual disclaimer applies.

1 Introduction

Natural gas transmission is a typical network industry. Theoretical (Sharkey 1982) and empirical evidence (Gordon et al. 2003) underline the subadditivity in the cost structure and therefore gas transmission companies remain highly regulated. The purpose of this paper is to provide empirical evidence of a robust benchmarking technique for regulation when the number of regulated companies and/or data observations is small.

Since the late 1980s a substantial reform process was undertaken with the objectives of cost reductions and efficiency increases in regulated network industries. The transition from cost-plus regulation, where companies recover their costs with a fixed rate of return (Joskow 2006; Farsi et al. 2007) to incentive-based regulation is the latest development towards more efficient production and cost reduction. In an incentive-based regulatory framework, price and revenue caps are set based on the RPI-X formula (Littlechild 1983; Beesley and Littlechild 1989) where the determination of the expected efficiency savings (X) is usually based on empirical results obtained from sophisticated efficiency analysis approaches (also called benchmarking analyses). This framework, where the efficiency performance of the companies is evaluated against a reference performance (Farsi et al. 2005), has mainly been favored by European regulators and played a crucial role in the regulatory processes in the UK and the Nordic countries.

Using benchmarking methods in regulatory practice has been widely criticized (Shuttleworth 2005, 2003). One of the major criticisms is the low number of observations in this sector, for a robust and consistent benchmarking. As shown in Table 1 the low number of observations is caused by strong concentration and absence of competition in natural gas transmission (for Germany see e.g. Hirschhausen et al. (2007)). In fact, in most of the European countries, e.g., Finland and Belgium, a single transmission company is operating. In others, e.g., Spain, Sweden and Austria, several independent companies are operating. In Germany, for the first round of determining efficiency scores data on only 8 companies (due to legislation) are considered in the benchmarking procedure. Moreover, the regulator often collects data on a yearly basis, thus additionally restricting sample size. Hence, both the low number of companies and the yearly data basis severely limit sample size for a serious application of traditional benchmarking methods. Regulators of natural gas transmission system operators require guidance in adapting their models to the empirical challenges.

A possible solution to expand the number of observations is to use data from other countries found in international benchmarking exercises. The study CEER (2006) analyzes relative efficiencies of European natural gas transmission operators. Its sample consists of four European countries (one company each, covering different time spans (3-5 years)) and 43 US companies over 9 years. However, two major problems with international comparisons are the strong heterogeneity of firms and the differences in data definitions across countries. Europe has the added problems that data collection still remains a responsibility of national regulators and a harmonized consistent European data pool is not yet implemented. Hence, efforts are predominantly undertaken to establish national efficiency standards with a limited data sample that consolidate theoretical requirements and practical applicability.

A wide range of benchmarking approaches and frameworks exist in the literature (Jamasb and Pollitt 2001, 2003; Farsi et al. 2007) and the approaches can be separated into two main streams: nonparametric and parametric methods. Data Envelopment Analysis

(DEA) as a nonparametric approach and Stochastic Frontier Analysis (SFA) as a parametric framework are the most commonly used. The nonparametric methods determine the reference technology by means of linear programming methods whereas the parametric SFA assumes a functional relationship for the production process and determines the reference technology based on econometric methods. From a practical regulatory point of view both approaches have been useful to regulators: directly as part of the regulation process or as an additional control instrument for decision-making (Farsi et al. 2007). Both methods differ in their requirements for the underlying data volume in order to derive meaningful results.[1] Even if DEA, in terms of statistical properties, is more inefficient practical experience shows that DEA is used more frequently than SFA in the practical applications of efficiency analysis in the energy sectors, see Haney and Pollitt (2009); CEER (2006).

A further empirical challenge is that in regulatory practice a detailed benchmarking model, describing the production process by means of exact input and output variables of the firms is indispensable. Hence, the model should include as much relevant information as possible. This requires a reasonable number of observations to distinguish companies and derive meaningful results. However, given a pre-determined sample size, an increase in dimensions (i.e. more explanatory variables)—which might contribute to more appropriately modeling of reality—leads to fewer observations determining the efficiency frontier and therefore, to less information used to build it. This subsequently affects efficiency scores in nonparametric efficiency analysis. For example, utility regulation is often conducted on a yearly basis, making it impossible to increase sample size when all possible installations are already included in the sample. Hence, this practical obstacle often constrains the regulator's ability to meet the statistical requirements. However, reducing dimensions and conserving all available information at the same time improves the estimation of technical efficiency in a DEA framework.

A feasible solution is the application of principal components analysis (PCA) in DEA that reduces dimensions of the original set of variables whilst maintaining the information on variation of data (Haerdle and Simar 2003). The combination of DEA and PCA was proposed by Ueda and Hoshiai (1997), and Adler and Golany (2001, 2002) who aim to overcome the issue of over-estimation of relative efficiency due to large numbers of variables in DEA. They show that PCA can improve discriminatory power in DEA and give more reliable efficiency measurement in small samples. Fields of application refer mainly to network industries. Whereas Ueda and Hoshiai (1997) apply their approach to the telecommunication sector, Adler and Golany (2001) and Adler and Berechman (2001) refer to the airline industry, and Adler and Golany (2002) to university departments. Adler and Yazhemsky (2009) provide further theoretical developments and show the applicability of PCA to radial DEA models when only additive DEA models[2] were previously considered.

There are also other discrimination-improving approaches related to DEA. For example, Adler and Yazhemsky (2009) compare PCA with the approach of variable reduction based on partial covariance and find better performance of PCA. Podinovski and Thanassoulis (2007) controvert simple approaches, i.e. increasing the number of units and reducing the number of variables by means of aggregation or reduction, and more sophisticated ap-

[1]Simar and Wilson (2008) prove that the theoretical foundations of DEA are based on large datasets to produce meaningful results. By contrast, parametric approaches reveal a desirable feature in terms of consistency of the estimator, i.e. its convergence to the unknown parameter at a certain rate when sample size increases to infinity.

[2]For the difference between radial and additive models see Cooper et al. (2007).

proaches, where the latter can be grouped using additional information and additional measurements respectively. Additional measurements obtained by means of further treatment of data have an advantage over additional information. They do not require information that is not directly given by the data and that is often difficult to determine. Frequently, regulators are unable to identify more realistic profiles of an optimal mix of inputs and outputs that could be implemented in DEA by weight restrictions (Podinovski and Thanassoulis 2007). Weight restrictions modify the efficient boundary of the production possibility set such that unrealistic input-output-compositions are no longer used as reference. However, the PCA-DEA formulation causes similar effects without the need of additional information (Adler and Yazhemsky 2009) and is therefore also preferred for our empirical analysis.

This paper provides the first PCA-DEA (in terms of radial efficiency measurement) in the context of natural gas transmission regulation. Since European is not easily comparable, we use the US natural gas market as our reference model. The US natural gas market often serves as a reference model given the long and good record of regulatory experience and publicly available company data over the last three decades. Rather than potentially including US data in a European benchmarking exercise we use data on US natural gas transmission companies to illustrate how data limitations affect radial efficiency measurement and how PCA-DEA improves it. For a discussion comparing the US and European natural gas market see Jamasb et al. (2008). Our contribution to the literature and practical application is to support a pragmatic approach for European regulators who predominantly undertake efforts for national benchmarking and therefore face problems of limited data.

The remainder of the paper is structured as follows. Section 2 introduces traditional DEA methodology and describes the issue of small samples in nonparametric benchmarking. DEA is extended by means of PCA following Adler and Yazhemsky (2009). The model specifications are outlined in Section 3, which also presents the data we use. Within this section outlier detection is reviewed. Our results are presented in Section 4 and Section 5 concludes.

2 Methodology

DEA is a nonparametric method frequently used in regulatory practice to evaluate relative efficiency and to set company-individual efficiency targets subsequently. The reference technology is not determined by imposing a functional form that describes the production process or cost structure, but by piecewise linear programming assuming a transformation of inputs into outputs. However, basic DEA models consider two types of technology: constant returns to scale (CRS) proposed by Charnes et al. (1978), and variable returns to scale (VRS) suggested by Banker et al. (1984). The first translates into strict regulation practice assuming one optimal firm size whereas the latter allows for scale inefficiencies. We limit ourselves to assume VRS technology because it seems to be more reasonable in small samples (Adler and Yazhemsky 2009). We also impose input-orientation, meaning that input is minimized while output remains fixed. This is a reasonable and common assumption in network industries because firms are generally required to supply service to a fixed geographical area, and hence, the output vector is essentially fixed (Coelli and

Walding 2006, p. 59).[3]

The standard radial DEA environment incorporating VRS technology and minimizing individual relative efficiency θ can be written as the following linear program:

$$
\begin{aligned}
& \min_{\theta,\lambda} \ \theta \\
& \text{s.t. } Y\lambda - s_Y = Y_j \\
& \ -X\lambda - s_X = \theta X_j \\
& e\lambda = 1 \\
& \lambda, \theta, s_Y, s_X \geq 0
\end{aligned}
\tag{1}
$$

where θ represents the relative efficiency (that is the absolute efficiency of the unit under consideration relative to a maximum value of obtained efficiency by any of the units considered) of each company contained in the set $J = \{1, 2, \ldots, n\}$. X_i and Y_i are column vectors of k inputs and l outputs of unit j. Collecting the column vectors yields in a $k \times n$ matrix for inputs X and a $l \times n$ matrix for outputs Y respectively. The input and output weights are given by the column vector λ. The constraint $e\lambda = 1$ ensures that the VRS-restriction is taken into account.[4] The slack variables s_X and s_Y permit the optimization problem to be of linear form. Furthermore, s_X, s_Y, λ, and θ are supposed to be nonnegative.

To obtain meaningful results with DEA the number of relevant input and output variables must be in proportion to the number of observations. Regulatory practice demands a sophisticated model with a high number of inputs and outputs to describe the production process or cost structure realistically. How well the method is able to sufficiently discriminate between utilities becomes an issue particularly when the data are limited, which is a known issue in real regulatory practice. This is addressed by principal components analysis (PCA), which can be used to reduce dimensions (number of variables) of the optimization problem by means of constructing linear combinations of the original data (Adler and Golany 2001, 2002). This conversion alters the original coordinate system (Adler and Yazhemsky 2009). Selecting the number of linear combinations then can reduce the dimensions of the new coordinate system. The number of dimensions comprising this new coordinate system depends on satisfying a selection criterion, e.g., the Kaiser-Guttmann-criterion or the Joliffe-criterion. We follow the study by Adler and Golany (2002) who select two as the number of principal components that satisfy discrimination purposes. However, we exclusively consider the limitation of dimensions in terms of outputs since there is no problem with a single input. Thus, for the purpose of translating output data, the correlation matrix C is obtained from the output matrix Y with $Y = [Y_1, Y_2, \ldots, Y_l]$. The eigenvectors v_l given by C are used to create linear combinations of the form $PC_{Y_i} = \sum_l (v_{l,i} \times Y_l)$ that are also known as principal components (PC). Each of the principal components explains a certain ratio of the original variables' variance, whereby this ratio corresponds to the eigenvalues of C η_l. Commonly, eigenvalues are in descending order, and so are therefore principal components, i.e. PC_1 covers most of the variation in the data, PC_2 covers less of it, and PC_l covers the lowest proportion.

[3] Input-orientation can be implemented in parametric and nonparametric approaches. For a parametric application see for example Farsi et al. (2005).

[4] Relaxing this constraint yields CRS technology, i.e. $\lambda \geq 0$.

Here we consider the combination of PCA and radial DEA models according to Adler and Yazhemsky (2009). However, one drawback of PCA-DEA is its requirement of data transformation. In PCA-DEA data are transformed initially by PCA and have to be remodeled to the original form after optimization. It appears that only some radial DEA settings are tolerant towards data transformation. Pastor (1996) proves output translation invariance for input-oriented DEA models under VRS assumption. Hence, in general, the optimal solution using original data does not change when data are transformed. Although, translation invariance is not supported by all DEA models, their general properties are not affected by PCA-DEA, see Adler and Yazhemsky (2009).

For one unmodified input and all outputs to be transformed into principal components, the dual linear programming under VRS assumption can be written as follows:

$$
\begin{aligned}
&\min_{\theta,\lambda}\ \theta \\
&\text{s.t. } Y_{PC}\lambda - L_Y s_{PC} = Y_{PC,j} \\
&-X\lambda - s_X = \theta X_j \\
&L_Y^{-1} Y_{PC} \geq s_{PC} \\
&e\lambda = 1 \\
&\lambda, \theta, s_{PC}, s_X \geq 0
\end{aligned}
\tag{2}
$$

where $Y = [y_1, y_2, ..., y_p]$ is the matrix of p outputs and x the single input vector we use. L_Y is the matrix collecting the output weights obtained by PCA. The original data are weighted and enter through principal components Y_{PC} where $Y_{PC} = l_i^t Y = l_{1i}y_1 + l_{2i}y_2 + l_{3i}y_3 + l_{4i}y_4$ and l_i are the normalized eigenvectors from the correlation matrix of Y. Because all outputs are transformed into principal components, the minimization problem does not include separate output vectors. Both the slack variable s_{PC} and the original output data are weighted by the linear coefficients obtained by PCA.[5] As stated in formulation (1) VRS technology and nonnegativity of parameters and slack variables are assumed. If and only if all PCs are included, i.e. PCs explain 100% of the original data variation, the solutions of formulation (1) and (2) are equivalent (Adler and Yazhemsky 2009).

3 Model specification and data

3.1 Model specification

We want to determine the pipelines' relative ability (pipelines refer to companies operating such facilities) to provide services at least cost where we consider the demand as fixed in the short-term. Hence, the model set up is based on the idea of a cost driver analysis, meaning that costs are explained by output variables that are relevant to costs of the pipelines under consideration. This approach deviates from the purely technical representation of the production process by physical data but is often applied in regulatory practice, see e.g., CEER (2006) and Bundesnetzagentur (2006). An important issue that arises almost immediately

[5]Due to data transformation a new constraint enters the linear problem which ensures the slack variable to be equal or smaller than the product of inverse weighting matrix and weighted output data.

when applying benchmarking in regulatory practice, is cost comparability. There are essentially two ways of constructing the benchmarking basis, i.e. the short-run maintenance model and the long-run service model. For a broad discussion see Burns et al. (2005). The first model incorporates operating expenditures while the second model incorporates total expenditures (operating expenditures plus capital costs). Although the total cost approach offers some advantages, the evaluation of capital costs still must be conducted carefully and in a reliable manner. However, in practice regulators more often rely on the first model (Haney and Pollitt 2009), and therefore, we conduct our analysis of efficiency on the basis of the short-run maintenance model. The determination of variables to be included is discussed broadly in the literature. A comprehensive investigation of the variables to use as cost measures and cost drivers for international benchmarking and regulation purposes is presented by CEER (2006); Jamasb et al. (2008) examine the productivity development of US natural gas transmission companies and review the literature with respect to variables. We note that most of the studies presented in the latter paper rely exclusively on parametric approaches.

We develop two model settings (Model 1 and Model 2), each containing the same cost measurement but differ in their number of cost drivers. We select total operating and maintenance expenses (OPEX) as the input to be minimized.[6] Although there are arguments in favor of total expenses including capital costs, we do not consider them here. However, CEER (2006) shows high correlation between these two measurements. The basic model (Model 1) treats total amount of natural gas delivered (TotDeliv), transmission system (TransSys), peak deliveries (PeakDeliv), and total installed horsepower of compressor stations (HorPow) as OPEX determinants and therefore outputs. The second model (Model 2) adds transmission system losses (TransLos), which is an undesired output and therefore must be treated differently. It is not our aim to present the particular effect of this undesired output itself; rather, we wish to demonstrate how an additional output will affect the empirical analysis and therefore regulatory consequences.

For the purpose of demonstration and comparison, each of the two models is specified under traditional DEA and PCA-DEA methodology, both assuming VRS technology. The resulting four model specifications are listed in Table 2.

3.2 Data

We use data on US American natural gas transmission companies. The US natural gas industry offers a comprehensive record of publicly available data and regulatory history, making it ideal for our analysis. We compile data from the US Federal Energy Regulatory Commission's (FERC) database of the major interstate natural gas pipelines. This covers each natural gas company whose combined gas transported or stored for a fee exceed 50 million dekatherms in each of the previous three calendar years (FERC 2008, p. i). In total our original sample contains 37 US American natural gas transmission companies in 2007 operating only onshore pipelines.[7] However, these companies are either stand alone units or units covering a broader business portfolio (holdings). Table 3 summarizes all variables we use.

[6]This is known as OPEX-benchmarking. Haney and Pollitt (2009) list international regulators who in fact conduct OPEX regulation.

[7]We omit companies which also operate offshore pipelines since the technology differs.

The sample includes natural gas transmission pipelines that spend about 2,860 million USD on operating and maintenance for approximately 127,783 miles of onshore facilities. This covers about 66.5% of total US interstate pipeline mileage. Pipelines differ in transmission system[8] and total deliveries[9], ranging from 49.93 million Dekatherms (Dth) to over 6,046 million Dth. The data indicates that some deliver low amounts of gas in peak times[10] with a minimum of 0.19 million Dth, while others deliver up to the maximum 8.44 million Dth. Another output is compressor stations' total installed horsepower, an important characteristic of gas transmission. Installed horsepower (Hp) is calculated as the product of the number of stations and their certified horsepower. This enables us to incorporate a capacity measurement. In fact, the data show significant differences in installed horsepower ranging from a minimum of 9 million Hp to a maximum of nearly 1,435 million Hp. The standard deviation of 371.72 million Hp indicates the strong variation in the data.

An additional output variable is transmission system losses. In total nearly 39.7 million Dth of natural gas are lost that would not occur in total deliveries. Pipelines report data ranging from no losses to 6,685 thousand Dth. A record of zero losses is technically very unlikely. Therefore, we suspect measurement errors, which we try to overcome with the subsequent outlier detection. TransLos must be treated differently from the others because of the inverse interpretation of undesirable outputs. To ensure a correct representation, we translate this variable such that more losses are disadvantageous to companies' performance. Thus, we subtract from a large number[11] and choose 10,000,000 as the large number.[12]

3.3 Outlier detection based on super-efficiency

Because nonparametric methods are sensitive to outliers (Simar 2003), we conduct an outlier detection based on the concept of super-efficiency. Following Banker and Chang (2006), we choose the selection criterion of 1.2: companies achieving an efficiency score equal to or smaller than 1.2 are accepted for the sample and those exceeding this criterion are excluded from further analysis. We find that three of the 37 utilities are super-efficient: Columbia Gas Transmission Corporation with 5.42, Petal Gas Storage, L.L.C. with 2.97, and Vector Pipeline L.P. with 2.02. In addition, this outlier detection confirms doubts from reporting non-transmission system losses for two of the three.[13] Hence, our final sample size is 34 pipelines.

[8] Petal Gas Storage, L.L.C. operates the smallest pipeline system, 59 miles, and Tennessee Gas Pipeline Company operates the largest, 14,463 miles.

[9] Natural gas delivered does not only account for own sales but also for interactions with others.

[10] Natural gas delivered in peak times refers to single day peak deliveries summing deliveries to interstate pipelines and "others".

[11] Other ways to implement undesirable outputs in the DEA framework are discussed in Dyson et al. (2001).

[12] The results are insensitive to a variation of the large number to 8,000,000 instead.

[13] The other two of the four pipelines which report zero transmission losses do not determine the frontier in the super-efficiency analysis.

4 Results

This section presents our results.[14] First, we deal with the results of the PCA, followed by the efficiency estimation for the two models (Model 1 without TransLos and Model 2 with TransLos) and methodologies (DEA and PCA-DEA). We then discuss the results for our model specifications for a particular pipeline to illustrate the relevance of PCA-DEA for real-world regulatory practice.

4.1 Principal components analysis

The principal components analysis enables us to reduce the dimensions of the linear program and thus to increase discrimination between the pipelines of interest. Table 4 shows the results of our separate PCA analysis for both models. In terms of output, the first principal component (PC_1) captures at least 82% of data variation in both models. Considering also PC_2 results in a cumulative explanation of more than 95% in Model 1 and 90% in Model 2 of the total data variation. Using only these two output PCs does not cause much loss of information for either Model 1 (4.73%) or Model 2 (9.47%). Since we consider only one input PC, we capture all information. Hence, it exactly represents the single input and does not affect efficiency measurement.

4.2 Efficiency of pipelines

Descriptive statistics of the pipelines' individual efficiencies (by percentage) given by DEA and PCA-DEA for each model are shown in Table 5. A company is fully efficient if it achieves 100%. The lower the efficiency score the worse the company has performed relative to its peers. We find two general results. First, compared to the traditional DEA approach, PCA-DEA yields lower efficiency. For example, pipelines in Model 1 (without TransLos) achieve 66.89% on average but 46.54% under the PCA-DEA specification. This empirically reflects the argument of Adler and Yazhemsky (2009, p. 3) by which PCA-DEA has effects similar to the imposition of weight restrictions, which renders parts of the efficient boundary of the production possibility set no longer efficient. In other words, companies that are really specialists in one of the original dimensions would be considered efficient performers due to linear programming. In fact, only specialization in this particular dimension would lead to efficiency, whereas the overall performance of the affected company does not. The single feature criterion (specialist in one dimension) is a particular problem for nonparametric approaches, while the weights of the variables by the coefficients attenuate the empirical problem in parametric SFA frameworks (Riechmann and Rodgarkia-Dara 2006). This overestimation of efficiency occurs especially when only a few observations are present relative to the number of variables. By means of PCA-DEA we reduce the space to only two dimensions and thus improve the efficiency determination.

Second, comparing the particular specifications of Model 1 with their counterparts in Model 2 (including TransLos), we observe higher efficiency in the latter model. This observation is almost true for every statistic, e.g., DEA specification in Model 1 reveals a mean of 66.89% and 77.55% in Model 2, and PCA-DEA specification reveals a mean of 46.54% in Model 1 and 60.04% in Model 2. We observe only one exception in the minimum efficiency

[14]For calculations we use the PCA-DEA Program developed by Adler (`http://pluto.huji.ac.il/\~msnic/PCADEA.htm`).

scores of PCA-DEA specification. So far, both models appear to differ in some respects, e.g., to median or 75%-quantile scores, which is highly relevant to regulatory practice. However, the robustness of PCA-DEA analysis is supported when considering pipeline-specific efficiency scores.

Figure 1 shows how company-specific efficiency scores change with DEA and PCA-DEA, and with our two model specifications. In addition to the findings already discussed—also retraceable here—other noticeable findings occur. In both graphs the pipelines are arranged in increasing order of total deliveries (TotDeliv), indicating their size.

For DEA specification, none of the plots suggests an identifiable trend of better performance depending on pipelines' size. This can be explained by the VRS approach. However, for PCA-DEA specification, the larger pipelines seem to be better performers. Intuitively, the impact of single features, which make companies efficient in the range of smaller companies when VRS technology is assumed, is attenuated.

The number of pipelines that are part of the efficiency frontier is clearly higher when DEA applies. In this case, Model 1 depicts seven efficient utilities, and Model 2 even defines half of the sample as efficient due to the additional output variable TransLos. It is for technical reasons that the more variables are included in traditional DEA, the more units are considered to be efficient. This has particular importance in small samples. Moreover, Adler and Yazhemsky (2009) show by means of Monte Carlo simulation, that a trade-off occurs between incorrect classification of (in)efficient decision-making units under traditional DEA and PCA-DEA. If technology and salient variables are correctly specified, traditional DEA never defines truly efficient units incorrectly as inefficient, i.e. the probability of error type 1 is zero. But at the same time, the probability of incorrectly defining inefficient units as efficient (error type 2) is high in DEA under VRS. Thus, we can expect a remarkable proportion of pipelines to be over-estimated in terms of efficiency, and potential cost reduction would remain uncovered. Therefore, the aim of regulation is not achieved.

While PCA-DEA can improve benchmarking activities while notably lowering the level of over-estimation, there is a cost. PCA-DEA causes a certain level of under-estimation. However, in radial efficiency measurement, this effect is minor. Adler and Yazhemsky (2009) demonstrate that with PCA-DEA the probability of under-estimation (error type 1) is very small while the probability of over-estimation (error type 2) significantly improves. Empirically PCA-DEA in our analysis defines three (Model 1) and nine (Model 2) pipelines as efficient. Note that in both cases only two PCs are included in the analysis and thus, the ratio of variables and observations is acceptable. Hence, PCA-DEA offers methodological features that are preferable to those of traditional DEA.

However, for both models we observe that most of the pipelines suffer from introducing PCA-DEA. In Model 1, the second-smallest pipeline delivering about 53 million Dth of natural gas (MIGC, LLC) achieves 50.79% under the DEA specification and decreases to 37.66% under PCA-DEA; a larger pipeline delivering 1,360 million Dth of natural gas (Dominion Transmission, Inc.) achieves 69.84% under DEA and decreases to 58.22% under PCA-DEA. But in Model 1 there are also companies that do not suffer from introducing PCA-DEA, i.e. those delivering 50 (Guardian Pipeline, LLC), 421 (IPOC as Agent/Iroquois Gas Trans. Sys. LP), and 3,270 (Transcontinental Gas Pipe Line Corporation) million Dth. We note that only peers (fully efficient companies) remain at the same level as before. It seems that their respective efficiency score is not distorted from unique characteristics[15]

[15] Riechmann and Rodgarkia-Dara (2006) point out that statistical fuzziness and unique characteristics

and full efficiency is justified.

According to Adler and Yazhemsky (2009, p. 10) it is preferable to avoid the omission of relevant variables because it leads to under-estimation of the mean efficiency. For regulatory practice including operating characteristics, quality variables, etc., in a sophisticated model can be important. However, the request for a realistic representation of company structures easily increases the number of variables substantially and hence, harms the ratio between observations and variables. The known consequence is a deteriorated discrimination capability of DEA. In fact, including TransLos in place of the mentioned variables yields significantly changed efficiency scores in both model specifications of Model 2, i.e. DEA and PCA-DEA. Still, the methodological difference induces a reduction of dimensions when PCA-DEA is applied; thus, using PCA-DEA does not affect the discriminatory capability although more variables are considered before. When we compare the PCA-DEA results of Model 1 (without TransLos) and Model 2, 29% of the companies (10 out of 34) exhibit lower efficiency under Model 2. Other companies improve or remain as good as before. The maximum individual worsening of 3.93% is experienced by the company delivering 100 million Dth in total (Equitrans LP). Note that because dimensions are equal in both models, changes seem to be associated with new information. At the same time, the PCA-DEA specification in Model 2 discloses the ability of PCA to account for specialists which we explain by one specific pipeline in more detail in the following section.

4.3 Case Study: Northern Border Pipeline Company

Northern Border Pipeline Company delivers 907 million Dth in total and achieves very low efficiency scores in Model 1 (27.02% with DEA and 19.23% with PCA-DEA), but the efficiency scores increase significantly when including TransLos in Model 2. Under traditional DEA, the pipeline achieves 100% efficiency. This indicates specialization in the particular variable TransLos which accounts for roughly 78 thousand Dth (so it seems unlikely to be an error in reporting). In contrast, when applying PCA-DEA, the efficiency score falls to 80.72%. What cannot be seen from this graph directly is how the reference set of Northern Border Pipeline Company changes between Model 1 and Model 2 with respect to PCA-DEA specification.

Table 6 provides more insight on the relevance of this reference set (peers) on the efficiency of our example. In Model 1 Northern Border Pipeline Company is compared to the efficient utilities IPOC as Agent/Iroquois Gas Trans. Sys. LP and Transcontinental Gas Pipeline Corporation, whereas in Model 2 Dominion Transmission, Inc. and El Paso Natural Gas Company appear to be its peers.[16] Obviously, IPOC as Agent/Iroquois Gas Trans. Sys. LP is a much smaller company, e.g., OPEX are only 9.3 mn USD, and total deliveries account for 420.6 mn Dth. Peers in the reference set of Model 2 are structurally more alike than the peers in Model 1. This can also be observed in Figure 1, where in Model 1 (PCA-DEA specification) the peers of Northern Border Pipeline Company are the efficient companies delivering 421 and 3,270 mn Dth, and in Model 2 the peers are those efficient pipelines delivering 1,360 and 6,047 mn Dth. This finding confirms the idea of DEA in the regulatory context. Burns et al. (2005, p. 304) relate benchmarking techniques

are sources of distortion.

[16]The peers in Model 1 with DEA specification are Transcontinental Gas Pipeline Corporation and Guardian Pipeline, L.L.C. In Model 2 with DEA specification Northern Border Pipeline Company serves as the peer, because of its specialization.

to yardstick competition and point out that one key feature of DEA is that it identifies "local" conditions, i.e. analyses the efficiency of a firm with reference to other firms that are similar in their combinations of outputs, for example. If regulators want benchmarking to fulfill this prerequisite, our results support its fulfillment when relevant variables are part of the analysis and discrimination power is given by applying PCA-DEA.

5 Conclusion

The purpose of this paper is to empirically demonstrate how improving discriminatory power in nonparametric efficiency analysis affects the efficiency scores of natural gas transmission companies. Moreover, we desire to support a pragmatic approach of efficiency evaluation for (European) regulatory authorities that accounts for a poor ratio between the number of variables and the number of observations.

Over the last decades network industries with natural monopoly character have experienced extensive restructuring towards incentive-based regulation schemes. Restructuring aims to motivate more efficient production and cost structures. Benchmarking has become an established tool in regulatory practice to identify company-individual targets for achieving these goals. Although there is an increasing interest in parametric benchmarking methods, e.g., SFA, practical experience show frequent application of nonparametric approaches such as DEA. For meaningful efficiency measurement, DEA requires a sufficient amount of data. However, due to the former monopolistic market structures and yearly conducted efficiency evaluation, this cannot always be guaranteed in reality. Limited data negatively affects DEA's discriminatory power, and thus increases the probability of efficiency over-estimation. This issue amplifies when a large number of variables are considered to describe the production process or cost structure of companies. To address this issue, DEA can be combined with PCA. By means of linear combinations of the original variables PCA reduces the dimensions while maintaining a large proportion of the variation in the original data. Consequently, discriminatory power in PCA-DEA improves and results in more robust efficiency scores. If regulators want benchmarking to fulfill this prerequisite, our results support its fulfillment when relevant variables are part of the analysis and discrimination power is given, i.e. by applying PCA-DEA. We test our hypotheses by applying PCA-DEA to a large sample of US natural gas transmission pipelines. We chose to employ US data because it is publicly available and the industry has a significant regulatory record. We defined two models, one with four output variables and a second with five; both models had a single input.

Our results suggest that PCA-DEA improves nonparametric efficiency analysis. Models applying traditional DEA display a high proportion of fully efficient pipelines (up to 50%), where we can suspect many are over-estimated. Because over-estimation decreases, pipelines on average perform less well under PCA-DEA than under DEA, which we trace back to more realistic efficiency measurement. We then show that additional outputs significantly change the results and, in PCA-DEA models, improve the evaluation of pipelines. Efficiency score changes between the different PCA-DEA model specifications appear to be not due to higher model dimensions, but due to worthwhile information and structurally similar reference companies. We conclude that these findings support current regulatory practice by mitigating the conflict between too few observations, and the demand for many variables to produce an appropriate representation of the relevant structures.

References

Adler, N., and Berechman, J. (2001). Measuring airport quality from the airlines's viewpoint. An application of data envelopment analysis. *Transport Policy*, *8*, 171-181.

Adler, N., and Golany, B. (2001). Evaluation of deregulated airline networks using Data Envelopment Analysis combined with Principle Component Analysis with an application to Western Europe. *European Journal of Operational Research*, *132*, 260-273.

Adler, N., and Golany, B. (2002). Including principle component weights to improve discrimination in Data Envelopment Analysis. *Journal of Operational Research Society*, *53*, 985-991.

Adler, N., and Yazhemsky, E. (2009). Improving discrimination in Data Envelopment Analysis: PCA-DEA or Variable Reduction. *European Journal of Operational Research*, *doi: 10.1016/j.ejor.2009.03.050*.

Banker, R. D., and Chang, H. (2006). The super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research*, *175*, 1311-1320.

Banker, R. D., Charnes, A., and Cooper, W. (1984). Some models for estimating technical and scale inefficiency in Data Envelopment Analysis. *Management Science*, *30*, 1078-1092.

Beesley, M., and Littlechild, S. (1989). The regulation of privatized monopolies in the United Kingdom. *Rand Journal of Economics*, *20*(3), 454-472.

Bundesnetzagentur. (2006). Bericht der Bundesnetzagentur nach § 112a EnWG zur Einfuehrung der Anreizregulierung nach § 21a EnWG. `http://www.bundesnetzagentur.de/media/archive/6715.pdf`. Accessed 2 November 2009.

Burns, P., Jetkins, C., and Riechmann, C. (2005). The role of benchmarking for yardstick competition. *Utility Policy*, *13*.

CEER. (2006). International benchmarking and regulation of European gas transmission utilities. `http://www.energy-regulators.eu`. Accessed 5 May 2009.

Charnes, A., Cooper, W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. *Journal of Operational Research*, *2*, 429-444.

Coelli, T., and Walding, S. (2006). Performance measurement in the Australian water supply industry: A preliminary analysis. In T. Coelli and D. Lawrence (Eds.), *Performance measurement and regulation of network utilities* (chap. 3). Cheltenham, UK: Edward Elgar Publishing.

Cooper, W. W., Seiford, L. M., and Tone, K. (2007). *Data Envelopment Analysis, a comprehensive text with models, applications, references and DEA-Solver software.* New York, USA: Springer.

Dyson, R., Allen, R., Camanho, A., Podinovski, V., Sarrico, C., and Shale, E. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*, *132*, 245-259.

Farsi, M., Fetz, A., and Filippini, M. (2007). *Benchmarking and regulation in the electricity distribution sector.* (CEPE Working Paper No. 54, ETH Zrich, Zurich, Switzerland)

Farsi, M., Filippini, M., and Greene, W. (2005). Efficiency measurement in network industries: Application to the Swiss railway companies. *Journal of Regulatory Economics*, *28*, 69-90.

FERC. (2008). FERC Financial Report. FERC Form No. 2: Annual Report of Major Natural Gas Companies and Supplemental Form 3-Q: Quarterly Financial Report. http://www.ferc.gov/docs-filing/forms/form-2/form-2.pdf. Accessed 12 May 2008.

Gordon, D., Gunsch, K., and Pawluk, C. V. (2003). A natural monopoly in natural gas transmission. *Energy Economics*, *25*, 473-485.

Haerdle, W., and Simar, L. (2003). *Applied multivariate statistical analysis.* Berlin, Germany: Springer.

Haney, A. B., and Pollitt, M. G. (2009). *Efficiency analysis of energy networks: An international survey of regulators* (EPRG Working Paper 0915, Cambridge Working Paper in Economics 0926). University of Cambridge, Electricity Policy Research Group.

Hirschhausen, C. von, Neumann, A., and Ruester, S. (2007). *Competition in natural gas transportation? Technical and economic fundamentals and application to Germany* (Report for EFET). Technical University Dresden and DIW Berlin.

Jamasb, T., and Pollitt, M. G. (2001). Benchmarking and regulation: International electricity experience. *Utilities Policy*, *9*(3), 107-130.

Jamasb, T., and Pollitt, M. G. (2003). International benchmarking and yardstick regulation: An application to European electricity distribution utilities. *Energy Policy*, *31*(15), 1609-1622.

Jamasb, T., Pollitt, M. G., and Triebs, T. (2008). Productivity and efficiency of US gas transmission companies: A European regulatory perspective. *Energy Policy*, *36*, 3398-3412.

Joskow, P. J. (2006). Regulation of natural monopolies. In A. M. Polinsky and S. Shavell (Eds.), *Handbook of Law and Economics.* Amsterdam: North-Holland.

Littlechild, S. (1983). *Regulation of British telecommunications profitability.* (Report to the Secretary of State, London)

Pastor, J. T. (1996). Translation invariance in Data Envelopment Analysis: A generalization. *Annals of Operations Research*, *66*, 93-102.

Podinovski, V. V., and Thanassoulis, E. (2007). Improving discrimination in Data Envelopment Analysis: Some practical suggestions. *Journal of Productivity Analysis*, *28*, 117-126.

Riechmann, C., and Rodgarkia-Dara, A. (2006). Regulatorisches Benchmarking - Konzeption und praktische Interpretation. *Zeitschrift fuer Energiewirtschaft*, *30*, 205-219.

Sharkey, W. W. (1982). *The theory of natural monopoly.* Cambridge: Cambridge University Press.

Shuttleworth, G. (2003). Firm-specific productive efficiency: A response. *The Electricity Journal*, *16*(3), 42-50.

Shuttleworth, G. (2005). Benchmarking of electricity networks: Practical problems with its use for regulation. *Utilities Policy*, *13*(3), 310-317.

Simar, L. (2003). Detecting outliers in frontier models: A simple approach. *Journal of Productivity Analysis*, *20*, 391-424.

Simar, L., and Wilson, P. W. (2008). Statistical inference in nonparametric frontier models: Recent developments and perspectives. In H. O. Fried, C. K. Lovell, and S. S. Schmidt (Eds.), *The Measurement of Productive Efficiency and Productivity Growth* (p. 421-521). New York, USA: Oxford University Press.

Ueda, T., and Hoshiai, Y. (1997). Application of Principle Component Analysis for parsimonious summarization of DEA inputs and/ or outputs. *Journal of the Operations Research, Society of Japan*, *40*(4), 466-478.

Tables

Table 1: European regulated TSO natural gas companies

Country	Number	Country	Number
Austria	7	Latvia	1
Belgium	1	Lithuania	1
Czech Republic	1	Luxembourg	1
Denmark	1	Netherlands	1
Estonia	1	Poland	1
Finland	1	Portugal	1
France	2	Romania	1
Germany	20	Slovakia	1
Greece	1	Slovenia	1
Hungary	1	Spain	8
Ireland	1	Sweden	3
Italy	2	UK	1

Source: Technical Annex to the Communication from the Commission to the Council and the European Parliament COM(2009)115.

Table 2: Model specification

	Model 1		Model 2	
	DEA	PCA-DEA	DEA	PCA-DEA
DEA	✓		✓	
PCA-DEA		✓		✓
VRS	✓	✓	✓	✓

Table 3: Descriptive statistics of US natural gas transmission companies, onshore (2007)

Variable	Opex	Total Deliveries	Transmission System	Peak Deliveries	Installed Horsepower	Transmission System Losses
Unit	mn USD	mn Dth	Miles	mn Dth	thou Hp	thou Dth
Sum	2,860.32	34,191.24	127,783.20	86.81	11,003.22	38,677.68
Minimum	1.25	49.93	59.00	0.19	9.00	0.00
Maximum	402.67	6,046.71	14,463.20	8.44	1,434.27	6,684
Mean	77.31	924.09	3,453.60	2.35	125.95	1,045.34
Median	31.50	403.89	1,680.40	1.68	297.38	615.66
Std. Dev.	99.61	1,255.53	3,703.33	2.12	371.72	1,399.32

Source: FERC Form No. 2.

Table 4: Principal components analysis for Models 1 and 2

Variance explained by principal component in %				
	Model 1		Model 2	
PC	Input	Output	Input	Output
1	100	87.76	100	82.19
2		7.52		8.34
3		3.35		5.71
4		1.38		2.68
5				1.08

Table 5: Efficiency of US American natural gas transmission companies in %

	Model 1 (without TransLos)		Model 2 (with TransLos)	
Statistic	DEA	PCA-DEA	DEA	PCA-DEA
Minimum	27.02	19.23	30.65	19.10
25%-quantile	44.78	31.46	53.83	39.52
Mean	66.89	46.54	77.55	60.04
Median	63.86	39.51	93.45	48.39
75%-quantile	95.53	57.08	100	98.50
Maximum	100	100	100	100

Table 6: Peers of Northern Border Pipeline Company in PCA-DEA model specifications

Variable	Opex	TotDeliv	TransSys	PeakDeliv	HorPow	TransLos
Unit	mn USD	mn Dth	Miles	mn Dth	thou Hp	thou Dth
NBPC	165.3	907.0	1,399	2.6	536.6	77.9
	Peers in Model 1					
I/I	9.3	420.6	414	1.4	78.3	489.4
TGPC	117.3	3,270.0	10,325	8.4	1,434.3	6,684.6
	Peers in Model 2					
DTI	70.7	1,360.1	3,344	4.0	350.2	398.5
EPNGC	373.4	6,046.7	10,240	5.1	1,136.4	3,038.8

Notes: NBPC = Northern Border Pipeline Company, I/I = IPOC as Agent/Iroquois Gas Trans. Sys. L.P., TGPC = Transcontinental Gas Pipeline Corporation, DTI = Dominion Transmission, Inc., EPNGC = El Paso Natural Gas Company

Source: FERC Form No. 2.

Figure 1: Pipeline-individual efficiency of US American natural gas transmission companies